Causal Effect of Social Services on Crime Reduction: A Study of Defunding the Police

Richard Li

Stats 209 (Fall 2020)

Contents

1 Introduction

This proposed project seeks to answer the question: is investing in social services more effective at improving public safety than funding police departments?

The purpose of this work is to inform evidence-based policymaking and city budgeting, in order to optimally allocate limited public resources to improve public safety with minimal harm to marginalized communities.

1.1 Social Disorder Hypothesis, DAGs

The underlying hypothesis is that most forms of crime arise due to social disorder – impoverishment, homelessness, drug addiction, mental illness, and unemployment. These interlocking disorders create scarcity and destabilize marginal communities, driving illegal markets and other criminalized behavior (Draine 2002).

Disinvestment from social services, driven by decades of neoliberal US policies, have magnified these social disorders (Wacquant 2009). Over the same period, the rise of mass incarceration has absorbed many people on the margins of society. The incarcerated population is disproportionately Black and poor (Bobo 2010), and people with mental illness are overrepresented (Fuller 2015).

Under this hypothesis, the expansion of the carceral system, for which police serve as the initial contact, serves to manage and control marginal populations, without addressing the underlying social disorders (Alexander 2010). Instead, investing in social programs (public housing, job training) and trained professionals (mental health specialists, social workers) may prove more effective at reducing crime in the long run, than investing in more police.

These proposed causal relationships are represented in a DAG in Figure 1. Note that the treatment variables (program funding) are all indirectly connected to the outcome (crime rate), via the social disorders. Also, we have bidirectional arrows despite being a DAG, as we believe the two causal effects can occur at different time periods.
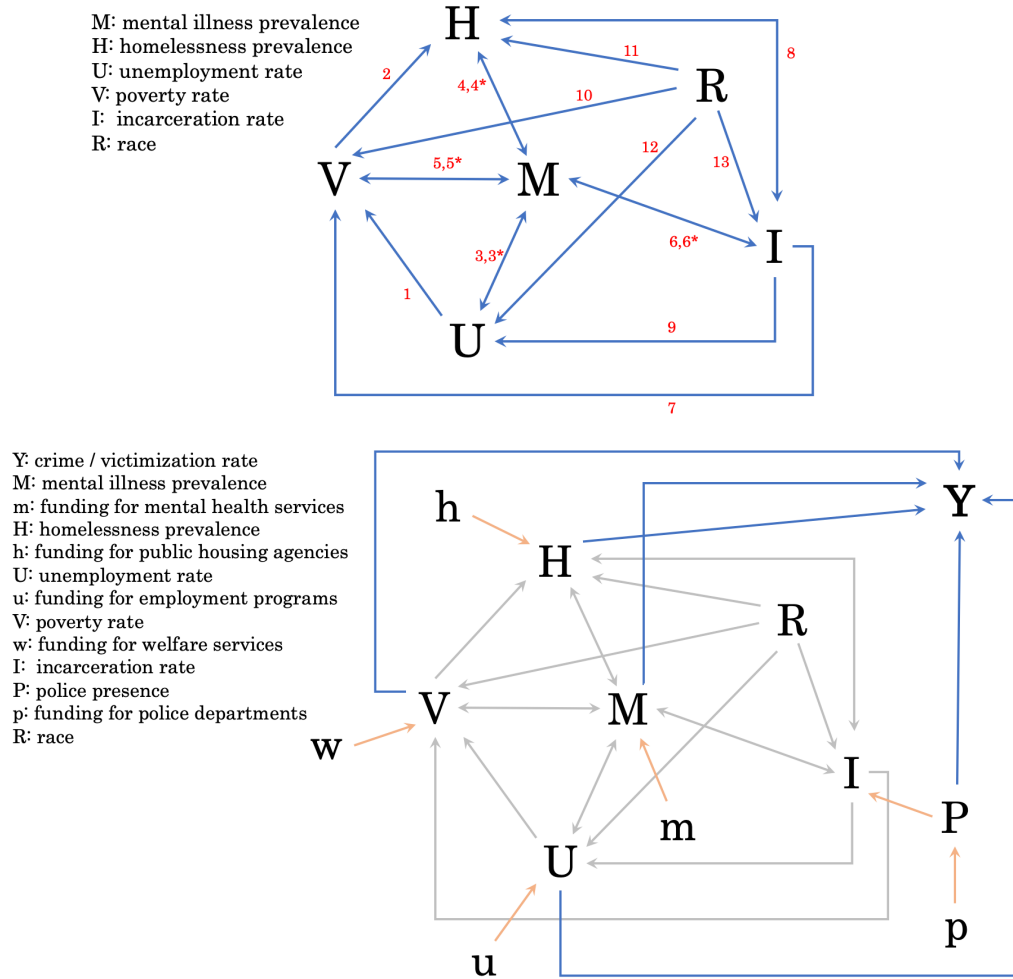
Figure 1: **Top**: Partial DAG showing proposed relationships between social disorder variables (upper case). **Bottom**: Full DAG, showing funding sources (treatments; lower case), and crime rates (outcome; Y). Note that all treatments are indirectly related to the outcome via the social disorder variables.

Discussion on relationships between social disorders (top): Intuitively, chronic unemployment leads to poverty (1), which increases risk of homelessness (2). Mental illness can make it difficult to maintain employment (3), undergo the arduous process of applying for public housing or Section 8 (4), and maintain financial solvency (5). People with visible mental illness are also disproportionately arrested and victims of police killings (6) (Fuller 2015). On the flip side, unemployment, homelessness, poverty, and incarceration are major risk factors for developing mental illness (3*, 4*, 5*, 6*). Formerly incarcerated individuals lose eligibility for public services, including welfare (7) and public housing assistance (8), and are allowed to be legally discriminated against in the private housing market (8) and job applications (9). Black families continue to be subject to mortgage lending discrimination (10), housing discrimination by private landlords (11), and employment discrimination (12) (Alexander 2010). Black communities are disproportionately surveilled by police and incarcerated (13) (Logan 2017).

1.2 Competing Theories on Criminality

A competing theory of criminality, supported by decades of criminological research into deviance, posits that criminal behavior arises from a rational decision-making process, where an individual weighs the risks and rewards of breaking the law, and chooses to engage in criminal behavior, in part because the risk of arrest or cost of guilty plea are too low (Akers 1990). This supports the expansion of police forces and more punitive laws.

This theory is supported by the success of hotspot policing, where concentrated police activity has been shown to reduce violent crime at that location and time (Braga 2005). However, these studies do not measure if crime was simply displaced spatially or temporally.

Further, there is no evidence that mass incarceration has led to crime reduction. Nonetheless, we acknowledge that police presence may have an effect on crime rates. Thus, we design our study to estimate the causal effects of 1) funding social programs on crime rates, and 2) funding police departments on crime rates, and then comparing the two.

1.3 Data Sources and Outcome Variable

We have 30 years of Census data on city budgets, including funding for police departments and various social programs, for 10-20 US cities ("Annual" 2019). For these cities and years, we also have data on crime rates, broken down by types of crime (NIBRS 2020).

One major challenge of this study is that police presence is a confounder – it inevitably leads to more discovery of crime, even if the actual crime rate has not changed. Ideally, we choose an outcome variable that is invariable to police presence, like victimization rates, which are reported by victims of crime and not police (NCVS 2019). Unfortunately, these data are not available at the granular level of cities / jurisdictions that we require.

Further, focusing exclusively on crime rates as the outcome variable or optimization objective overlooks the well-documented harms that overpolicing has on communities (Kochel 2011). These externalities – harrassment, discrimination, and improper use of force – are not given any weight when evaluating program success. If these societal harms were measured, quantified, and included in the cost function – along with reduction in crime rates – it may better represent the tradeoffs of increasing police activity on public safety. We hope to be able to incorporate this outcome in our analysis someday.

In summary, we would ideally want our outcome variable to be an overall measure of public safety, which accounts for both the adverse effects of crime on victims, as well as the effects of overpolicing on communities. In the absence of this data, we use crime rates as a proxy for public safety.

2 Assignments and Potential Outcomes

Our data is best described as panel data, as we observe sequential treatments (funding) and outcomes (crime rates) over time, for the same set of units (cities). Thus, our data is two dimensional – both cross-sectional and longitudinal.

Thus far in class, we have applied the potential outcomes framework to cross-sectional data, where many units are treated at one point in time. We now extend this potential outcomes framework to the panel setting.

The majority of the methods here are adapted from Bojinov et al 2020, which develops methods for causal inference on panel data that are very analogous to what we learned in class.

This formulation focuses on finite populations, where uncertainty is derived solely from the randomized assignment, instead of from sampling from a superpopulation. The superpopulation assumption might not be unreasonable in our application, as there are thousands of city and county jurisdictions with independent city budgets and police departments. Nonetheless, we proceed with this model-free, finite population approach.

Outline of paper: First, we define assignments and potential outcomes in the panel setting, and their relevant assumptions (Section 2). Then, we define a family of lag-p estimands (Section 3). Next, we provide an unbiased estimator, and derive its finite population asymptotic distribution (Section 4). This allows us to conduct Neymanian tests and Fisher randomization tests (Section 5). Finally, we discuss some considerations around covariate matching (Section 6). This paper primarily focuses on project scope and design, paving the way for actual analysis in the future.

2.1 Assignment Paths

In our panel, we have N units (cities), indexed by $i \in \{1, ..., N\}$, observed over T time periods, indexed by $t \in \{1, ..., T\}$. For each city i and period t, we allocate an assignment $Z_{i,t} \in \mathcal{Z}$. Tentatively, each assignment will be the relative change in funding from the previous year, for that social program (i.e. 1.2 or 0.8).

The assignment path is the series of assignments for unit i over the sample period, $Z_{i,1:T} = (Z_{i,1}, ... Z_{i,T})' \in \mathcal{Z}^T$. The assignment panel contains all the assignment paths for units N:

$$Z_{1:N,1:T} = \begin{pmatrix} Z'_{1,1:T} \\ \vdots \\ Z'_{N,1:T} \end{pmatrix} \in \mathcal{Z}^{N \times T} \tag{1}$$

2.2 Potential Outcome Panels

The potential outcome for a unit i at time t along assignment panel $z_{1:N,1:T}$ is $Y_{i,t}(z_{1:N,1:T})$. Most generally, the potential outcomes depends on the entire

assignment panel. However, we assume that there is no treatment spillovers across units (***non-interference***). This is a reasonable assumption as long as police from one city are not dispatched to another, or residents from one city do not benefit from social programs in another. This should be the case as long as our jurisdictions chosen are non-overlapping. Thus we can simplify $Y_{i,t}(z_{1:N,1:T})$ $= Y_{i,t}(z_{i,1:T})$.

The next assumption we make is ***non-anticipating treatments***, which claims that current potential outcomes are not influenced by future assignments. This should be reasonable, as cities typically have little ability to predict future funding decisions. This simplifies $Y_{i,t}(z_{i,1:T}) = Y_{i,t}(z_{i,1:t})$. The formal definition can be found in Bojinov 2020.

Finally, we assume ***perfect compliance***. This is reasonable as we do not expect social programs to refuse changes to their funding. For an observed assignment panel $w_{1:N,1:T}^{obs}$, the observed outcome panel is $y_{1:N,1:T}^{obs} = Y_{1:N,1:T}(w_{1:N,1:T}^{obs})$.

2.3 Assignment Mechanisms

We make several assumptions about the assignment mechanism, which are analogous to the cross-sectional case in class. These assumptions allow us to conduct different inference tests in Section 5.

The first is that assignments are ***sequentially randomized***, such that current assignments do not depend on future potential outcomes. This is reasonable as the city budgeting process as no knowledge of future crime rates. This is the analog of the "unconfounded" or "ignorable" assumption in the cross-sectional case, and can be written as:

$$\Pr(Z_{1:N,t}|Z_{1:N,1:t-1} = z_{1:N,1:t-1}, Y_{1:N,1:T}) = \Pr(Z_{1:N,t}|Z_{1:N,1:t-1} = z_{1:N,1:t-1}, Y_{1:N,1:t-1}(z_{1:N,1:t-1}))$$
$$(2)$$

Next, we assume that assignments are ***contemporaneously independent***, which means at time t, assignments are made independently across units. This does not limit the assignment at time t from depending on past observed outcomes and assignments *across all units.* For instance, after New York's stop-and-frisk program was hailed as a success, other cities may invest in similar police programs in the future. The formal definition is in Bojinov 2020.

A stronger version of this is the ***individualistic*** assumption, which requires that the assignment for unit i at time t is independent of past assignments and outcomes *for all other units.* This would not hold perfectly (as described in the example above), but jurisdictions do behave independently for the most part. I would suspect that this assumption is most violated between cities close in geography or political stance.

Finally, we assume the assignment mechanism is ***probabilistic***, meaning the probability of seeing any assignment z is between 0 and 1. I do not anticipate that funding changes are ever entirely deterministic.

3 Estimands

3.1 Unit-Level Estimands

We first define some families of estimands that can be applied to a potential outcome panel for a single unit i at time t. The most general estimand we could use is the *dynamic causal effect*, which compares the potential outcomes along entire assignment paths $z_{i,1:t}, \tilde{z}_{i,1:t} \in \mathcal{Z}^t$.

$$\tau_{i,t}(z_{i,1:t}, \tilde{z}_{i,1:t}) := Y_{i,t}(z_{i,1:t}) - Y_{i,t}(\tilde{z}_{i,1:t}) \tag{3}$$

We focus on one subset of estimands, called the *lag-p dynamic causal effect*, which we abbreviate LDCE. The LDCE compares the outcomes of two assignment paths that diverge only after time $t - p$, for $p + 1$ periods. Prior to time $t - p$, the assignments are fixed to the observed assignment path (depicted in Figure 2).

$$\tau_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) := Y_{i,t}(z_{i,1:t-p-1}^{obs}, \mathbf{z}) - Y_{i,t}(z_{i,1:t-p-1}^{obs}, \tilde{\mathbf{z}}) \tag{4}$$
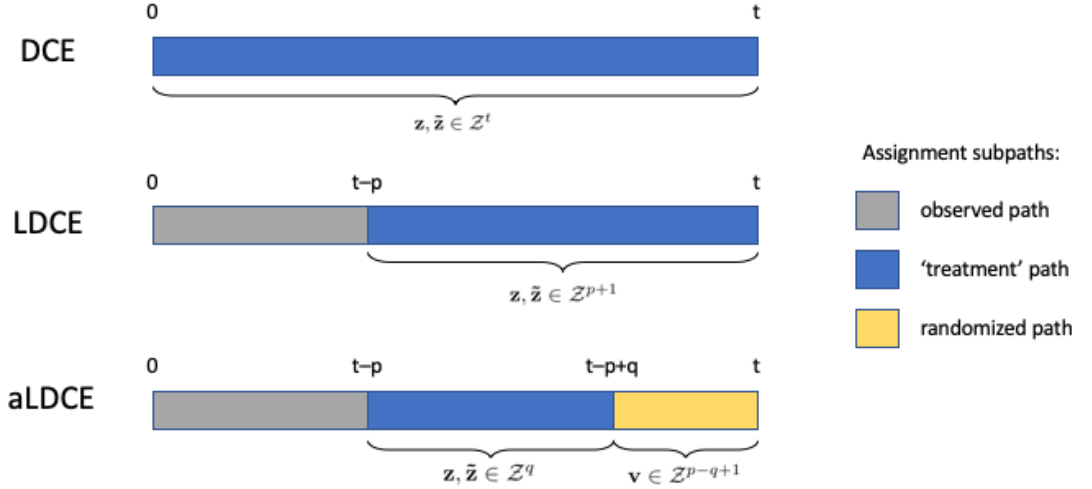

Figure 2: Graphical depiction of different unit-level estimands. The general *dynamic causal effect* (DCE) compares potential outcomes along entire assignment paths of length t. The *lag-p dynamic causal effect* (LDCE) compares the potential outcomes for two assignment paths that share the same observed path up to time $t - p$, after which they differ. The *weighted average lag-p,q dynamic causal effect* (aLDCE) summarizes the effect of switching assignment paths from period $t - p$ to $t - p + q$ on outcomes at time t, by averaging over all possible assignment paths afterwards.

A variation of the LDCE is the *weighted average lag-p,q dynamic causal effect* (aLDCE), which further restricts the divergent assignment paths to q periods, from time $t - p$ to $t - p + q$ (Figure 2), after which all possible assignment paths are averaged (Figure 2).

$$\tau_{i,t}^{\dagger}(\mathbf{z}, \tilde{\mathbf{z}}; p, q) := \sum_{\mathbf{v} \in \mathcal{Z}^{p-q+1}} a_{\mathbf{v}} [Y_{i,t}(z_{i,1:t-p-1}^{obs}, \mathbf{z}, \mathbf{v}) - Y_{i,t}(z_{i,1:t-p-1}^{obs}, \tilde{\mathbf{z}}, \mathbf{v})] \quad (5)$$

This aLDCE reduces the dependence of the observed treatment path on the definition of the estimand. The weights $a_{\mathbf{v}}$ are selected to reflect the probability of various assignment paths after the q-region of interest. In the following sections, we build upon the LDCE estimand, although the aLDCE can be directly substituted in all cases.

3.2   Average Estimands

Next, we average these unit-level estimands over units and time periods, in order to leverage the information

The *time-t lag-p average dynamic causal effect* (t-LaDCE[1]) averages over all units at time t:

$$\bar{\tau}_{.t}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{1}{N} \sum_{i=1}^{N} \tau_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \quad (6)$$

The *unit-i lag-p average dynamic causal effect* (i-LaDCE) averages over all time periods for unit i (depicted in Figure 3):

$$\bar{\tau}_{i.}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{1}{T - p} \sum_{t=p+1}^{T} \tau_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \quad (7)$$

The *total lag-p average dynamic causal effect* (total LaDCE) averages over all units and all time periods:

$$\bar{\tau}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{1}{N(T - p)} \sum_{t=p+1}^{T} \sum_{i=1}^{N} \tau_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \quad (8)$$

The total LaDCE is valuable, as it allows us to identify the effect of funding changes on crime rates $p = 5$ years later (for instance), over all N cities and $T = 30$ years. Figure 3 explains how this temporal averaging is done.

The t-LaDCE can also be helpful to compare how treatment effects evolve over time. For instance, social programs may be less effective during periods of economic recession, when people are subject to more threats to their stability. Alternatively, i-LaDCE can be used to evaluate heterogeneity within units. For instance, we can evaluate if social programs are more effective in certain cities than others.

---

[1] Note that the LaDCE is distinct from the unit-level aLDCE defined above.
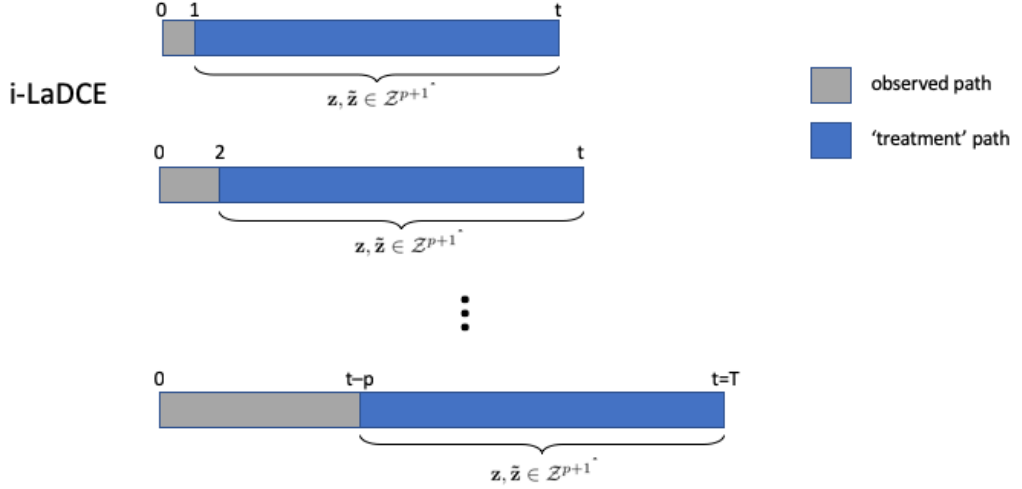
Figure 3: Graphical depiction of how temporal averaging is done in the *unit-i lag-p average dynamic causal effect* (i-LaDCE). The i-LaDCE considers all subpaths up to $t = T$, with lag of length p. For instance, for $p = 5$, it considers the outcomes at year 6 due to treatment since year 1, the outcomes at year 7 due to treatment since year 2, etc. Therefore, we use all the information in our length T time series, with treatment windows of length p.

3.3 Applicability to our Study

3.3.1 Selection of lag-p

For our study, using these lag-p estimands means that we are measuring the impact of funding changes from up to 5 years ago (for instance), on crime rates today. Importantly, this lag-p value must be predetermined and cannot be changed dynamically. This is a challenge, as it is very unclear how long it takes for funding changes to be reflected in the reduction of social disorders. It likely varies by type of social service.

For instance, mental health and addiction counseling often require years of treatment, characterized by periods of relapse and remission. This may require a longer timescale to see tangible changes in public safety, compared to deploying more police to displace or incarcerate those with mental health disorders.

Further, increases in funding may require more time to materialize in tangible outcomes than cutting funding. It takes significantly more time to hire new employees and construct new programs, than it does to fire people and shut down programs.

I anticipate that I will experiment with different lag-p values, potentially informed by correlation strengths or literature. If I have distinct lag-p values for each assignment variable $\mathbf{z}$ and for each estimand, I will have to justify these decisions based on the context of each research question.

3.3.2 Selection of treatment paths

Another key design decision is what we will choose the 'treatment' paths $\mathbf{z}$ and $\tilde{\mathbf{z}}$ to be. Most simply, we could focus on treatment paths of length 1 ($q = 1$ for aLDCE), and discretize our observed treatments into 3 categories - increase funding, decrease funding, or stay the same: $\{+, -, 0\} \in \mathcal{Z}$.

Then we could compare how increased funding affects outcomes compared to constant funding: $\mathbf{z} = +$ and $\tilde{\mathbf{z}} = 0$. Alternatively, we could compare $\mathbf{z} = -$ and $\tilde{\mathbf{z}} = 0$, or $\mathbf{z} = +$ and $\tilde{\mathbf{z}} = -$.

There are two extensions we could make to this simple case. First, we could increase the length of treatment path (q), which would allow us to evaluate a series of funding decisions on the outcome. This is more realistic, as organizations typically require years of continuous investment to build programs. However, this complicates how we choose our treatment paths. We could look at the total treatment effect, where $\mathbf{z} = (1, ..., 1)$ and $\tilde{\mathbf{z}} = (0, ..., 0)$, but it is unlikely we will observe many instances of this in our data, especially for longer treatment paths.

Hypothetically, it might be nice to design treatment paths such as 'increasing for at least 2 out of 3 periods', or 'no funding decreases'. This would require a set of treatment paths, i.e. $\mathbf{z} \in \{(+, +, 0), (+, 0, +), (0, +, +)\}$. I am not sure if defining this type of treatment path is permissible for this estimand.

Another extension would be discretizing the treatments more granularly, to distinguish between big and small increases. This could help answer questions such as: is a funding increase for social programs more effective at reducing crime than a comparable funding increase for police? Again, this runs into the issue that there will be fewer observed instances of any particular assignment path, the more granular the assignments are.

I am also uncertain if maybe this type of question (where is \$1 best invested) is better answered through regression methods. Would it be correct to use causal inference methods with more crude assignments ($\{+, -, 0\} \in \mathcal{Z}$) to first prove there is a causal effect, and then use regression methods to demonstrate the strength of correlation?

4 Estimators

We first define the adapted propensity score as the probability of seeing treatment path $\mathbf{z}$ given the previous assignment path and all potential outcomes:

$$p_{i,t-p}(\mathbf{z}) := \Pr(Z_{i,t-p:t} = \mathbf{z} \mid Z_{i,1:t-p-1}, Y_{i,1:t}(Z_{i,1:t-p-1}, \mathbf{z})) \tag{9}$$

Since we only observe the outcomes along the realized assignment path, we cannot calculate the above propensity score for any arbitrary path $\mathbf{z}$, but we can calculate $p_{i,t-p}(z_{i,t-p:t}^{obs})$.

4.1 Unit-Level Estimators

Recall the *lag-p dynamic causal effect* (LDCE) is $\tau_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) := Y_{i,t}(z_{i,1:t-p-1}^{obs}, \mathbf{z}) - Y_{i,t}(z_{i,1:t-p-1}^{obs}, \tilde{\mathbf{z}})$. We define a Horvitz-Thompson estimator, with an analogous structure to the cross-sectional version from class:

$$\hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{Y_{i,t}(z_{i,1:t-p-1}^{obs}, \mathbf{z})\, \mathbb{1}(z_{i,t-p:t}^{obs} = \mathbf{z})}{p_{i,t-p}(\mathbf{z})} - \frac{Y_{i,t}(z_{i,1:t-p-1}^{obs}, \tilde{\mathbf{z}})\, \mathbb{1}(z_{i,t-p:t}^{obs} = \tilde{\mathbf{z}})}{p_{i,t-p}(\tilde{\mathbf{z}})} \tag{10}$$

We can simplify the numerator by noting that we have either the observed outcome or 0:

$$\hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{y_{i,t}^{obs}\, [\mathbb{1}(z_{i,t-p:t}^{obs} = \mathbf{z}) - \mathbb{1}(z_{i,t-p:t}^{obs} = \tilde{\mathbf{z}})]}{p_{i,t-p}(z_{i,t-p:t}^{obs})} \tag{11}$$

This HT estimator is computable. Without showing the proof here (Bojinov 2020), we note that this estimator is unbiased, as long as the assignment mechanism is individualistic and probabilistic (Section 2):

$$\mathbb{E}[\hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \mid \mathcal{F}_{i,t-p-1}] = \tau_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \tag{12}$$

Similarly without proof, Bojinov finds that the upper bound on the variance can be estimated by:

$$\widehat{Var}(\hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \mid \mathcal{F}_{i,t-p-1}) = \frac{(y_{i,t}^{obs})^2\, [\mathbb{1}(z_{i,t-p:t}^{obs} = \mathbf{z}) + \mathbb{1}(z_{i,t-p:t}^{obs} = \tilde{\mathbf{z}})]}{p_{i,t-p}(z_{i,t-p:t}^{obs})^2} \tag{13}$$

4.2 Average Estimators

We can now average this unit-level estimator over time and/or units, analogous to how we calculated the average estimands above:

$$\hat{\bar{\tau}}_{\cdot,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{1}{N} \sum_{i=1}^{N} \hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \tag{14}$$

$$\hat{\bar{\tau}}_{i\cdot}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \tag{15}$$

$$\hat{\bar{\tau}}(\mathbf{z}, \tilde{\mathbf{z}}; p) := \frac{1}{N(T-p)} \sum_{t=p+1}^{T} \sum_{i=1}^{N} \hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \tag{16}$$

The unit-level estimator (Equation 11) follows the martingale difference property, which means the estimation error is conditionally independent over time and units.

This ensures that our average estimators are also unbiased for their respective estimands. The upper bounds on their variances and calculated similarly (Bojinov 2020):

$$\hat{\sigma}_{\cdot t}^{2} := \sum_{i=1}^{N} \widehat{Var}(\hat{\bar{\tau}}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \mid \mathcal{F}_{i,t-p-1}) \tag{17}$$

$$\hat{\sigma}_{i\cdot}^{2} := \sum_{t=p+1}^{T} \widehat{Var}(\hat{\bar{\tau}}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \mid \mathcal{F}_{i,t-p-1}) \tag{18}$$

$$\hat{\sigma}^{2} := \sum_{i=1}^{N} \sum_{t=p+1}^{T} \widehat{Var}(\hat{\bar{\tau}}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \mid \mathcal{F}_{i,t-p-1}) \tag{19}$$

4.3 CLT

With an unbiased estimator for the average estimands, and a conservative estimator for its variance, we can now conduct Neymanian inference using the finite population central limit theorem:

$$\frac{\sqrt{N} \left[\hat{\bar{\tau}}_{\cdot t}(\mathbf{z}, \tilde{\mathbf{z}}; p) - \bar{\tau}_{\cdot t}(\mathbf{z}, \tilde{\mathbf{z}}; p) \right]}{\sigma_{\cdot t}} \xrightarrow{d} N(0, 1) \quad as \ N \to \infty \tag{20}$$

$$\frac{\sqrt{T-p} \left[\hat{\bar{\tau}}_{i\cdot}(\mathbf{z}, \tilde{\mathbf{z}}; p) - \bar{\tau}_{i\cdot}(\mathbf{z}, \tilde{\mathbf{z}}; p) \right]}{\sigma_{i\cdot}} \xrightarrow{d} N(0, 1) \quad as \ T \to \infty \tag{21}$$

$$\frac{\sqrt{N(T-p)} \left[\hat{\bar{\tau}}(\mathbf{z}, \tilde{\mathbf{z}}; p) - \bar{\tau}(\mathbf{z}, \tilde{\mathbf{z}}; p) \right]}{\sigma} \xrightarrow{d} N(0, 1) \quad as \ NT \to \infty \tag{22}$$

Note that these estimators approach asymptotic distributions as either their units or time periods (or both) increase.

5 Inference

5.1 Neymanian Inference

With the unbiased average estimators and the central limit theorem presented above, we can conduct conservative inference for the average estimands $\bar{\tau}_{\cdot t}(\mathbf{z}, \tilde{\mathbf{z}}; p)$, $\bar{\tau}_{i\cdot}(\mathbf{z}, \tilde{\mathbf{z}}; p)$, and $\bar{\tau}(\mathbf{z}, \tilde{\mathbf{z}}; p)$.

We can conduct hypothesis testing of test weak nulls that the average dynamic causal effects are zero. For instance, to test $H_0 : \bar{\tau}_{\cdot t}(\mathbf{z}, \tilde{\mathbf{z}}; p) = 0$, we would calculate $\hat{\bar{\tau}}_{\cdot t}(\mathbf{z}, \tilde{\mathbf{z}}; p)$ and $\hat{\sigma}_{\cdot t}^2$ from the observed data, calculate the Z value, compare to the Z score for the chosen α value, and reject the null as appropriate.

Similarly, we could calculate confidence intervals and see if zero lies within the confidence interval.

5.2 Fisher Randomization Test

We can also conduct Fisher randomization tests, under the assumptions that the assignment mechanism is probabilistic and sequentially randomized (Section 2). We follow the same procedure as in class.

The FRT assumes a sharp null, where there is no treatment effect between any two assignment paths for any unit, i.e. $H_0 = \bar{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p) = 0$ for all $\mathbf{z}, \tilde{\mathbf{z}}$, and i. This allows us to fill in the science table by setting all potential outcomes $Y_{i,t}(w_{1:t-p-1}^{obs}, w)$ to the observed outcome $y_{i,t}^{obs}$, for all treatment paths w and all units i.

Next we carry out a simulation to randomly assign treatment paths, and then calculate the $\hat{\tau}_{i,t}(\mathbf{z}, \tilde{\mathbf{z}}; p)$ corresponding to each simulation. This provides a distribution of the test statistic under the sharp null, which we can compare with the observed value to to find a p-value or confidence interval.

6 Covariate Matching

In the case of observational studies, as we have, we cannot rely on randomization in the design to do inference. We need to make assumptions around *ignorability* (discussed in Section 2), which we test through sensitivity analyses.

Since we could not randomize across covariates and confounders, we need to perform either matching or subclassifcation/regression on these covariates.

When we test the causal effects between one treatment variable (i.e. mental health funding) and crime, we will treat all other treatment variables (funding for public housing, police, etc) as covariates on which we must match.

Unfortunately, observational studies and covariate matching are not discussed in the potential outcomes frameworks for either panels (Bojinov 2020) or time series (Bojinov and Shephard 2017). We discuss below how we could apply these concepts to our panel setting.

6.1 Propensity Score

One way to match units is on the propensity score, which would be defined as the probability of seeing assignment path $\mathbf{z}$, conditional on the covariates over that same time span ($t - p$ to t for LDCE):

$$\pi_{\mathbf{z}}(x_{i,t-p:t}) := \Pr(Z_{i,t-p:t} = \mathbf{z} \mid x_{i,t-p:t}) \tag{23}$$

where $x_{i,t-p:t}$ is a 2-dimensional matrix, across covariates and time periods. Since the covariates vary with time, we need to consider the covariate matching at all time steps.

One way of handling the added time dimension is to vectorize the covariate matrix to length Cp, where C is the number of covariates. This ensures that we regress on the covariates at all time steps:

$$logit(\pi_i) = \beta'_i x^{\text{1D}}_{i,t-p:t} + \epsilon_i \tag{24}$$

One weakness of propensity scores is the loss of information from reducing the dimensionality of all the covariates to a scalar. This procedure introduces 1 more dimensionality reduction, along time.

6.2 Multivariate Matching

Multivariate matching retains information from each covariate, instead of collapsing them into a single score. The benefits should be even more accentuated for our case where we have the added time dimension.

If we continue using the 1-D vectorized version as above, we can calculate the Mahalanobis (or its robust variant) without any adjustment:

$$d(x^{\text{1D}}_{i,t-p:t}, x^{\text{1D}}_{j,t-p:t}) = (x^{\text{1D}}_{i,t-p:t} - x^{\text{1D}}_{j,t-p:t})^T \hat{\Sigma}^{-1} (x^{\text{1D}}_{i,t-p:t} - x^{\text{1D}}_{j,t-p:t}) \tag{25}$$

Practically, this will increase our number of covariates by a factor of p, which could be significant. Presumably, better matches will be found for shorter lag-p periods, as there are fewer covariates to match on.

This 1-D vectorization approach may not work with the Mahalanobis distance, which detects correlated variables. It is unclear if a covariate may be correlated with itself at different time steps (autoregression?).

6.3 Matching Design

The choice of n in 1:n matching will likely depend on how many cities we have in the "treatment" group (however we define that) versus the "control". We would ideally like to include as many cities as possible in our inference tests.

We may choose to exact match on certain covariates that we need important (such as police funding). Depending on the distribution of matches, we may also use a propensity score caliper to improve poor quality matches.

6.4 Sensitivity Analysis

After matching, we will conduct inference and calculate a p-value as if we have a randomized design. We then need to conduct sensitivity analysis to evaluate 1) the deviation from ignorability and 2) errors in matching.

This is critical because we have reason to believe funding decisions are not individualistic between cities (see Section 2.3), and there are likely macroeconomic factors (like recessions) that affect all city budgets.

Further, our above method of vectorizing covariates across time may prove to be a poor decision for matching. Conducting sensitivity analysis allows us to see to what extent our assumptions are sensitive to violations up to Γ.

Guillaume – many thanks for helping me think through my project design. The papers you shared were very helpful. I look forward to carrying out this work in the next few months.

References

Akers, R. (1990). Rational Choice, Deterrence, and Social Learning Theory in Criminology: The Path Not Taken. The Journal of Criminal Law and Criminology (1973-), 81(3), 653-676. doi:10.2307/1143850

Alexander, M. (2010). The new Jim Crow: Mass incarceration in the age of colorblindness. The New Press.

"Annual Survey of State and Local Government Finances" (2019). US Census. https://www.census.gov/programs-surveys/gov-finances.html

Bobo, Lawrence D. and Victor Thompson. 2010. "Racialized Mass Incarceration: Poverty, Prejudice, and Punishment." in Doing Race: 21 Essays for the 21st Century, edited by Hazel R. Markus and Paula Moya. New York: Norton, 322-355.

Bojinov, Iavor Rambachan, Ashesh Shephard, Neil. (2020). Panel Experiments and Dynamic Causal Effects: A Finite Population Perspective.

Bojinov, Iavor Shephard, Neil. (2017). Time Series Experiments and Causal Estimands: Exact Randomization Tests and Trading. SSRN Electronic Journal. 10.2139/ssrn.2991825.

Braga, Anthony. (2005). Hot Spots Policing and Crime Prevention: A Systematic Review of Randomized Controlled Trials. Journal of Experimental Criminology. 1. 317-342. 10.1007/s11292-005-8133-z.

Draine, Jeffrey Salzer, Mark Culhane, Dennis Hadley, Trevor. (2002). Role of Social Disadvantage in Crime, Joblessness, and Homelessness Among Persons With Serious Mental Illness. Psychiatric services (Washington, D.C.). 53. 565-73. 10.1176/appi.ps.53.5.565.

Fuller, Doris A., et al (2015). The Role of Mental Illness in Fatal Law Enforcement Encounters. Treatment Advocacy Center.

Kochel, Tammy. (2011). Constructing Hot Spots Policing: Unexamined Consequences for Disadvantaged Populations and for Police Legitimacy. Criminal Justice Policy Review. 22. 350-374. 10.1177/0887403410376233.

Logan, J. R., Oakley, D. (2017). Black Lives and Policing: The Larger Context of Ghettoization. Journal of urban affairs, 39(8), 1031–1046.

"NCVS: National Crime Victimization Survey " (2019). Bureau of Justice Statistics. https://www.bjs.gov/index.cfm?ty=dcdetailiid=245

"NIBRS: National Incident-Based Reporting System" (2020). Federal Bureau of Investigation. https://www.fbi.gov/services/cjis/ucr/nibrs

Wacquant, L. (2009). Punishing the Poor: The Neoliberal Government of Social Insecurity. Duke University Press.